PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/67, 15/80, 15/81, C12P 21/02, C12N 15/55, 9/18, 9/36, 15/13, C07K 16/00

A2

(11) International Publication Number: **WO 00/05389**

(43) International Publication Date: 3 February 2000 (03.02.00)

(21) International Application Number: PCT/EP99/04919

(22) International Filing Date: 8 July 1999 (08.07.99)

(30) Priority Data:
98202432.5 20 July 1998 (20.07.98) EP

(71) Applicant *(for all designated States except AU BB CA CY GB GD GH IE IL KE LK LS MN MW NZ SD SG SZ TT UG ZA)*: UNILEVER N.V. [NL/NL]; Weena 455, NL–3013 AL Rotterdam (NL).

(71) Applicant *(for AU BB CA CY GB GD GH IE IL KE LK LS MN MW NZ SD SG SZ TT UG ZA only)*: UNILEVER PLC [GB/GB]; Unilever House, Blackfriars, London EC4P 4BQ (GB).

(72) Inventors: FRENKEN, Leon, Gerardus, Joseph; Unilever Research Colworth, Colworth House, Sharnbrook, Bedford, Bedfordshire MK44 1LQ (GB). SAGT, Cees; Universiteit van Utrecht, Heidelberglaan 8, NL–3584 CS Utrecht (NL). VERKLEIJ, Arie, J.; Universiteit van Utrecht, Heidelberglaan 8, NL–3584 CS Utrecht (NL). VERRIPS, Cornelis, Theodorus; Unilever Research Vlaardingen, Olivier van Noortlaan 120, NL–3133 AT Vlaardingen (NL).

(74) Agent: WURFBAIN, Gilles, L.; Unilever N.V., Patent Department, Olivier van Noortlaan 120, NL–3133 AT Vlaardingen (NL).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: PRODUCTION OF PROTEINS

(57) Abstract

A method for producing a heterologous protein in a lower eukaryotic cell, followed by secretion of said protein from said cell, said protein comprising at least one hydrophobic stretch, wherein the protein is glycosylated such as to increase the level of secretion of the protein, without substantially affecting the functionality of the protein and wherein at least one N–glycosylation group is located between the N–terminus of the protein and the first hydrophobic stretch.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Production of Proteins

Field of the invention

The present invention relates to the production of proteins, more particularly to a method for improving the secretion of proteins from eukaryotic cells.

Background of the invention

It is well known that the genetic modification of organisms can be used to produce proteins or fragments thereof. Generally, proteins which are produced in eukaryotic cells are initially located within intracellular organelles. For example proteins may initially be located in the endoplasmatic reticulum.

For many applications it is preferred that proteins, especially heterologous proteins, are adequately secreted extracellularly. For this purpose it is necessary that they can pass the cell plasma membrane in reasonable amounts and without substantial loss of protein activity.

A problem in the production of heterologous proteins in lower eukaryotic cells is that for some proteins, especially proteins with relatively large hydrophobic areas on their surface, the secretion is inefficient, in particular such proteins seem to have a low flux through the secretion system.

For example Frenken L.G.J. et al (1994) discloses that hydrophobic antibody fragments tend to accumulate in the endoplasmatic reticulum. Furthermore, Sagt, C.J.M. et al

(1998) discloses that in *Saccharomyces cerevisiae*, some cutinases with mutations to introduce hydrophobicity are secreted in significantly lower amounts than wild-type cutinase.

It has been suggested to improve secretion of heterologous proteins produced in a genetically modified organism by the introduction of a glycosilation site in the heterologous protein.

EP-A-704,527 relates to a process for the preparation of insulin. This document discloses that the introduction of N- glycosilation consensus site to a spacer region, which is not part of mature insulin molecule, remarkably increases expression in fungal cells, which are transformed with DNA sequences encoding such insulin precursors.

EP-A-394,951 discloses the introduction of a glycosilation site in fibroblast growth factor in order to induce secretion of this factor from transformed cells.

WO-A-96/05228, which is herewith incorporated by reference, relates to a method to produce single-chain Fv molecules in eukaryotic cells. This document discloses that glycosilation of single-chain Fv molecules can enhance the rate of secretion. For example an N-linked glycosilation site is introduced at position 19 in FR1 of V_H of single chain Fv molecule against human TfR (human transferrin receptor) and in FR1 of V_H of single chain Fv molecule against hapten DNP (Kurucz et al 1993) with an N linked glycosilation site at position 19 in FR1 of V_H. More specifically for example an N linked glycosilation site is introduced at position 19 of V_H of single chain Fv U7.6Ab-

sFv or OKT9 Ab-sFv. These proteins are excluded from the scope of the current invention.

It has been found however that glycosilation at an arbitrary position in the protein often does not lead to the desired increased secretion of a protein, especially not for secretion of proteins comprising a hydrophobic region.

Summary of the invention

Surprisingly applicants have found that especially good results are obtained if the location of the glycosylation group or groups is carefully chosen.
In particular it has been found that proteins with relatively large hydrophobic areas on their surface, which normally show a low secretion flux through the secretion system, can be secreted in significantly higher amounts if they are brought into glycosylated form, preferably N-glycosylated form, wherein the location of the glycosylation groups is chosen such that the hydrophobic stretches in the protein are shielded. On the other hand the location of the glycosylation groups is chosen such that they do not materially affect the functionality of the protein e.g. through influencing the binding affinity of the binding site and/or active site of the protein.

Accordingly in a first aspect the present invention relates to a method for producing a protein, preferably a heterologous protein, in a lower eukaryotic cell, followed by secretion of said protein from said cell, said protein comprising at least one hydrophobic stretch, wherein the

protein is glycosylated such as to increase the level of secretion of the protein, without substantially affecting the functionality of the protein and wherein at least one N-glycosylation group is located between the N-terminus of the protein and the first hydrophobic stretch and with the proviso that

a) the protein is not a single chain Fv molecule against human TfR with an N linked glycosilation site at position 19 in FR1 of V_H, and

b) the protein is not a single chain Fv molecule against hapten DNP with an N linked glycosilation site at position 19 in FR1 of V_H.

The invention is especially advantageous for improving the secretion of heterologous proteins. In a preferred embodiment of the invention, the protein can advantageously be de-glycosylated after secretion.

Detailed description of the invention

In the specification and claims the following terms and abbreviations are used.

As used herein, "eukaryotic cell" means a cell which comprises a nucleus containing the genetic material, surrounded by a cytoplasm which in turn is encompassed within the plasma membrane which marks the periphery of the cell.

A "gene" is a DNA sequence encoding a protein, including modified or synthetic DNA sequences or naturally occurring sequences encoding an RNA molecule, peptide, polypeptide, or protein and regions flanking the coding sequence involved in the regulation of expression.

A "hydrophobic stretch" of a protein is a sequence of amino acids in the protein that repel water. A quantitative definition of the term "hydrophobic stretch" is given in Blond-Elguindi, S. et al (1993). In this document a hydrophobic stretch can be identified by defining the Binding Protein (BiP) score of secreted proteins. This score consists of data generated with a peptide scan which determines the value of the BiP affinity for every 7 amino acids stretches of the protein. A value is given to the central amino acid of the 7 amino acids stretch. When the stretch contains hydrophobic amino acids at specific places the BiP is further increased. A total BiP score of 10 or more for a heptapeptide indicates a hydrophobic stretch. Hydrophobic amino acids are alanine, valine, leucine, isoleucine, proline, phenylalanine, tryptophan and tyrosine, methionine, and histidine. Preferred are valine, leucine, isoleucine, phenylalanine, tryptophan and methionine.

A "hydrophobically modified protein" is a protein which has an aggregated BiP score which is at least 50 units, more preferably at least 80 units, above the aggregated BiP score of the non-modified protein. The aggregated BiP score can be calculated by determining for each heptapeptide in the

protein an individual BiP score, in accordance to Blond-Elguindi, S. et al (1993). The aggregated BiP score can then be determined by calculating the sum of all individual BiP scores.

A "homologous protein" is a protein which is expressed in untransformed cells.

A "heterologous protein" is a protein which is not expressed in untransformed cells but which, under application of genetic engineering may be expressed in transformed cells.

A "ribosome" is a particle composed of ribosomal RNAs and ribosomal proteins that associate with messenger RNA and catalyse the synthesis of protein.

The "cytoplasm" is the contents of a cell that are contained within its plasma membrane but, in the case of eukaryotic cells, outside the nucleus.

The "cytosol" is the content of the main compartment of the cytoplasm, excluding the membrane bound organelles such as the endoplasmic reticulum and the mitochondria. Originall defined operationally as the cell fraction remaining after membranes, cytoskeletal components, and other organelles have been removed by low-speed centrifugation.

The "mRNA" (messenger RNA) is an RNA molecule that specifies the amino acid sequence of a protein. Produced by

RNA splicing from a larger RNA molecule made by RNA polymerase as a complementary copy of DNA. It is translated into protein by a process catalysed by ribosomes.

"Translation" is the process by which the sequence of nucleotides in a messenger RNA molecule directs the incorporation of amino acids into protein; occurs on a ribosome.

"N-terminus" (amino terminus) is the end of a polypeptide chain that carries a free alpha-amino group.

The present invention relates to an improved secretion of proteins, in particular heterologous proteins, from lower eukaryotic cells, said proteins comprising at least one hydrophobic stretch and/or said proteins being hydrophobically modified.

For the purpose of the invention the term protein is used to embrace proteins or fragments thereof, said fragments preferably comprising at least 10 amino acid residues, more preferred at least 50 amino acids, most preferred from 100 to 2,000 amino acids.

For the purpose of the invention, the protein is not a single chain Fv molecule against human TfR with an N linked glycosilation site at position 19 in FR1 of V_H, and the protein is not a single chain Fv molecule against hapten DNP with an N linked glycosilation site at position 19 in FR1 of V_H.

Eukaryotic cells for the purpose of the invention are lower eukaryotes such as for example moulds or yeasts. Preferred moulds belong to the genera *Aspergillus* or *Trichoderma*. Preferred yeasts belong to the genera *Saccharomyces*, *Kluyveromyces*, *Hansenula* or *Pichia*. Especially preferred is the use of host cells selected from *Saccharomyces cerevisiae*, *Aspergillus awamori* and *Pichia pastoris*.

The host cells are used for the production of the proteins, especially heterologous proteins. Well-known genetic manipulation techniques can be used to produce said heterologous proteins. For example a nucleotide sequence encoding the desired protein can be inserted into a suitable expression vector encoding the necessary elements for transcription and translation and in such manner that the protein will be expressed under appropriate conditions. The methods required to construct these expression vectors are well known to those skilled in the art. Preferred vectors stably integrate at pre-defined positions in the chromosome.

The present invention is particularly directed to improve the degree of secretion of proteins that normally have a low level of secretion. Another embodiment of the invention is especially preferred to be applied to the production of proteins which, under normal culturing conditions, are secreted at a level of less than 10 mg/g dry weight. Also the invention is especially advantageous for proteins which after glycosylation have a clearly improved level of

secretion, for example more than 1.5 times the secretion as compared to the non-glycosylated protein, more preferred more than 2.5 times, most preferably the level of secretion is improved by at least a factor of 3.5. Especially preferable, the level of secretion is more than 50 mg/g dry weight, more preferred more than 100 mg/g dry weight, most preferred more than 200 mg/g dry weight.

Also the invention is especially applicable to increase extracellular secretion of heterologous proteins which are not normally glycosylated. Preferably N-glycosylation is used. In order to introduce N-glycosylation during the production of a protein a acceptor group is needed which can either be Asn-X-Ser or Asn-X-Thr (where X is any amino acid except proline.

The molecular weight of the glycosylation groups is generally not critical.

The location of the glycosilation groups is chosen such that the hydrophobic stretches of the molecule are shielded e.g. to reduce the binding affinity of Binding Protein such as Bip to said hydrophobic stretches. On the other hand the location of the glycosilation groups is chosen such that they do not materially affect the functionality of the protein e.g. through influencing the binding affinity of the binding site and/or active site of the protein.

Both effects can be adequately predicted or can be measured by conventional binding tests. Applicants believe that it

will be well within the capability of the skilled person to select for each protein the appropriate size and location of glycosylation groups such as to increase the secretion of said protein.

In particular it is preferred that the glycosylation groups are orientated such that they provide a shielding effect for the hydrophobic areas of the protein when the protein is present in linear form in the Endoplasmatic Reticulum (ER) of the host cell. Additionally it is preferred that the glycosylation does not materially affect the binding and/or active site(s) of said protein when the protein is present in its folded three dimensional shape outside the production cell.

Although applicants do not wish to be bound by any theory it is believed that the beneficial effect of the glycosylation groups located at specific locations in the protein can be explained as follows.

In the production of proteins in lower eukaryotic cells, ribosomes are believed to bind mRNA molecules in the cytosol of the cell where translation of these nucleotide sequences will commence. Where a mRNA sequence encodes, at the 5 end, a signal peptide, the translation of this will cause a migration of the translating apparatus to the surface of the ER. The signal peptide then becomes embedded in the membrane of the ER and therefrom, protein synthesis by further translation of the mRNA will be concurrent with

the translocation of the developing protein molecule into the lumen of the ER.

Within the Endoplasmatic Reticulum, Binding Proteins such as BiP are present. These proteins tend to bind to hydrophobic stretches of the developing linear protein molecule. It is believed that this binding, which normally assists the folding of homologous proteins, leads to folding of protein molecules in such a way that their extracellular secretion from the host cell is much less efficient.

It is further believed that if a sequence of the mRNA encoding amino acids that will undergo glycosylation is translated prior to a sequence that encodes a hydrophobic stretch, this will result in an shielding effect being exerted on the hydrophobic stretch by the glycosylation group. This shielding is believed to reduce the binding affinity between binding proteins and the hydrophobic areas. By reducing the binding proteins that attach to the newly translated protein sequence the normal course of protein folding is prevented and thus the heterologous protein molecules are believed to retain a more linear conformation.

As indicated above, it is believed that the glycosylation groups are suitably present in part of the protein that is translated in the Endoplasmatic Reticulum prior to the hydrophobic stretch. To achieve this, the glycosylation site of the protein is located between the N-terminus of

the protein and the hydrophobic stretch of the protein. If more than one hydrophobic stretch is present in the protein, the glycosylation groups are preferably located between the N-terminus of the protein and the first hydrophobic stretch (i.e the hydrophobic stretch closest to the N-terminus). Even more preferably a glycosylation groups is located prior to each hydrophobic stretch i.e. the first glycosylation group between the N-terminus of the protein and the first hydrophobic stretch, the second glycosylation group between the first and the second hydrophobic stretch and so on.

In addition to adequate shielding of the protein in the Endoplasmatic Reticulum, it is also preferred that the location of the glycosylation groups is chosen such that they do not materially affect the functionality of the protein once it has been secreted.

Once secreted a protein normally is present in a specific three dimensional (3D) form. The functionality of a 3D structure of a protein is normally determined by the fact whether the binding sites and/or the active centres of the protein are inhibited in their functioning or not.

Applicants therefore believe that the location of the glycosylation groups in the protein should preferably be chosen such that the functionality of the binding site(s) and or the active centre(s) of the protein is not inhibited.

Applicants believe that this can be achieved by ensuring that the glycosylation groups are located at an adequate distance from said binding sites and/or active centres. In one embodiment this can be achieved by ensuring that the distance between the glycosylation sites and the binding site(s) and/or active site(s) is greater than the length of the glycosylation group. For example if the size of the glycosylation group is about 10 Angstrom, said distance should be at least 10 Angstrom, more preferred more than 12 Angstrom or even more than 15 Angstrom.

In another embodiment the distance between the binding sites and/or the active sites on the one hand and the glycosylation site at the other side can be less than the size of the glycosylation group. Under those circumstances, however, the orientation of the glycosylation should be chosen such that it cannot interact with the binding sites and/or the active sites. For example if the active site is located at the inside of a 3D protein, the glycosylation site may for example be located at the outer surface of the protein at a distance of less than the size of the glycosylation group to the active site, provided that the glycosylation group is at least partially directed away from the active site. For example said glycosylation group may point away from the surface of the protein and hence avoid interaction with the active site.

In a very advantageous embodiment the invention is applied to proteins having a generally roundish, egg- or ellipsoidal shape. Examples of such proteins are for

example lysozymes, cutinases and antibodies, especially variable fragments of antibodies. In these roundish shapes generally one binding area is present at one end of the ellipsoid (hereafter referred to as the upper end). The active site often is present near the binding site e.g. located at the inside of ellipsoid close to the binding site.

Applicants have now found that in elliposoidal proteins, preferably the glycosylation groups are located at the lower end of the ellipsoid. Preferably the distance between the glycosylation site and a hypothetical plane separating the binding site from the rest of the protein is more than 10 Angstrom, more preferred more than 12 or even more than 15 Angstrom.

Applicants believe that based on the above teaching it will be well within the ability of the skilled person to design the appropriate location of the glycosylation groups. For illustrating the invention more clearly, however the principles of the invention will be illustrated based on a number of actual embodiments.

Embodiment 1: cutinases

WO 94/14963 discloses various mutants of cutinase. As can be seen from figure 11 of this document the hydrophobic mutations of the cutinase are preferably located in a 15 Angstrom band around the Cα-117 plane. Also it is known that the active site of cutinase is located above the

hydrophobic band in the Cα-120 direction, i.e. at the top of the ellipsoid. Based on the above teaching it will be clear to the skilled person that in order to maintain the functionality of the protein after glycosylation, the preferred location for glycosylation is the area below the above described hydrophobic band in the Cα-115 direction i.e. the bottom of the ellipsoid. These modifications of the protein are illustrated in the examples.

Additionally the location of the glycosylation group should preferably be chosen such that strength of the BiP binding to the linear protein is reduced in the Endoplasmatic reticulum. As explained above this can advantageously be achieved by introducing a glycosylation group between the N-terminus of the protein and the first hydrophobic stretch.

For example a well-known hydrophobically modified cutinase is CY028 as described in C.M.J. Sagt et al, (1998). This enzyme has been made hydrophobic by the following mutations: G82A, A85F, V184I, A185L, L189F. This results in the introduction of two hydrophobic stretches, and hence the formation of a hydrophobically modified protein, having an aggregated BiP score which is about 80 units above the BiP score for the corresponding wild-type cutinase. As will be shown in the examples the introduction of these two hydrophobic stretches indeed results in a significant decrease in secretion efficiency.

To restore the secretion efficiency in accordance with the present invention, the skilled person would hence try to introduce a glycosylation site which satisfies the following criteria:

1. The glycosylation site is located between the N-terminus of the protein and the first hydrophobic stretch.
1. In the 3D structure the distance between the glycosylation site and the binding and/or active sites is sufficient to avoid interaction of the glycosylation group with said binding and/or active sites.

The 3D structure of this protein is given in figure 1. Figure 1 shows the 3D structure of cutinase CY028. As can be seen, the protein is generally ellipsoidal in shape. The longest axis of the representation corresponds to about 40-45 Angstrom. The binding domain (1) is located at the upper end of the representation, the active site (3) is located at the inside of the protein near the binding domain. Amino acid 29 is located at (2).

Suitably therefore the glycosylation groups can be located at the lower end of the ellipsoid e.g. below plane (4) which is located about 10 Angstrom downwards from the active site.

It is well-known in the art that to introduce glycosylation during the production of a protein an acceptor group is

needed which can either be Asn-X-Ser or Asn-X-Thr (where X is any amino acid except proline).

Therefore in order to ensure that adequate glycosylation takes place in accordance to the invention, the skilled person will seek possible mutations for the cutinase, which result in the availability of an additional acceptor group which can act as glycosylation site, said acceptor group being located before the first hydrophobic stretch and said receptor group in the 3D structure being located below plane(4) of figure 1.

The amino acid sequence of the cutinase gene is given in WO 94/14963 in figure 1D. From this it can easily be worked out that the A29S mutation (numbering starts at begin of the pro-sequence) leads to the introduction of a glycosylation site. This glycosylation site glycosylation in accordance to the invention, since this is both located before the first hydrophobic stretch and also located at the lower end of the 3D structure (evidenced by (2) in figure 1).

Embodiment 2: antibody (fragments)

Already at a very early stage during evolution, antibodies have been developed to protect the host organisms against invading molecules or organisms. Most likely one of the earliest forms of antibodies must have been developed in *Agnatha*. In these primitive fishes, antibodies of the IgM type consisting of heavy and lights chains have been detected. Also in many other forms of life ranging from amphibians to mammals, antibodies are characterized by the feature that they consist of two heavy and two light chains, although the heavy chains of the various classes of immunoglobulins are quite different. These heavy and light chains interact with each other by a number of different physical forces, but interactions between hydrophobic patches present on both the heavy and light chain are always important. The interaction between heavy and light chains exposes the complementarity determining regions (CDRs) of both chains in such a way that the immunoglobulin can bind the antigen optimally.

Heavy and light chains are composed of constant and variable domains. In the organisms producing immunoglobulins in their natural state the constant domains are very important for a number of functions, but for many applications of antibodies in industrial processes and products their variable domains are sufficient. Consequently many methods have been described to produce antibody fragments, but none of these have been particularly successful for large scale production at low costs.

European patent application EP-A1 584 431 (C. Casterman and R. Hamers) discloses the isolation of animal-derived immunoglobulins devoid of light chains (also indicated as heavy chain immunoglobulins), which can especially originate from animals of the camelid family (*Camelidae*). These heavy chain immunoglobulins are characterized in that they comprise two heavy polypeptide chains sufficient for the formation of one or more complete antigen binding sites, whereby a complete antigen binding site means a site which alone will allow the recognition and complete binding of an antigen, which can be verified by any known method regarding the testing of the binding affinity. The European patent specification further discloses methods for isolating these heavy chain immunoglobulins from the serum of *Camelidae* and details of the chemical structure of these heavy chain immunoglobulins. It also indicates that these heavy chain immunoglobulins and derivatives thereof can be made by using recombinant DNA technology in both prokaryotes and eukaryotes. Due to the absence of light chains in most of the immunoglobulins of *Camelidae* linkers are not necessary. The majority of the protein A-binding immunoglobulins of *Camelidae* consists just of two heavy chains and that these heavy chains are quite different from common forms of heavy chains, as the C_H1 domain is replaced by a long or short hinge.

The above illustrates that antibodies are proteins of high potential importance for industrial application. However, many of these applications cannot be realised because of the high costs of antibodies if produced by culturing cells.

Applicants have found that the present inventions can advantageously be applied to improve the secretion of various types of antibodies and fragments thereof. Especially the invention can advantageously improve the secretion of antibody fragments containing one or more variable domains. Examples of suitable antibody fragments are ScFv, V_H, V_L, and V_{HH}.

L.G.J. Frenken et al (1994) describes the production of ScFv, V_H and V_L Antibody fragments in *Saccharomyces cerevisiae*. It was found that antibody fragments such as ScFv-LYS, V_H and V_L can be produced in yeast but tend to accumulate in the Endoplasmatic Reticulum and hence have a low secretion efficiency. It is postulated that the secretion might be hampered by the formation of large aggregates, due to illegitimate interactions of the hydrophobic regions on the V_H and V_L fragments.

Applicants have now found that the secretion of antibodies or antibody fragments, in particular ScFv fragments or V_H and V_L fragments can markedly be improved if the antibody (fragments) are glycosylated at specific locations.

As described above the location of the glycosylation groups is chosen such that

1. in the linear form of the protein the glycosylation site is located between the N-terminus of the protein and a hydrophobic stretch of the protein; and

(b) in the 3D form the glycosylation groups are located such that they do not inhibit the functionality of the antibody fragments.

With respect to the location of the hydrophobic stretch in the linear protein, Example III describes the linear amino acid sequence of antibody fragments R2, R5 and R7 and replacement mutations thereof. In accordance to the invention the first requirement is that the glycosylation site is located between the N-terminus and the first hydrophobic stretch.

The second requirement is that the location of the glycosylation site is chosen such that the functionality of the antibody (fragment) is not materially affected. In this respect reference is made to figure 2 which shows the 3D structure of the R2 antibody fragment (production described in example III).

In the 3D structure it can be seen that the 3 CDRs (1) as indicated in black are located at the upper end of the antibody fragment structure. Also it can be seen that a plane (2) substantially passing through amino acids 25, 102, 94, 34, 51 and 57 (notation in accordance to Kabatt) separates the CDR area from the framework of the antibody fragment. Hence in accordance to the invention it is preferred that any glycosylation group in the 3D structure either has a distance to plane (2) which is greater than the size of the glycosylation group. For example if the size of the glycosylation group is about 10 Angstrom, the distance should be at least 10 Angstrom, more preferred more than 12 Angstrom or even more than 15 Angstrom. Alternatively the distance to plane (2) is smaller than the

size of the glycosylation group, but then the glycosylation group at least partially points away from said plane in order to avoid interaction of the glycosylation group with the CDRs (1).

As described above, in order to introduce glycosylation during the production of a protein an acceptor group is needed which can either be Asn-X-Ser or Asn-X-Thr or possibly Asn-X-Cys (where X is any amino acid except proline).

Therefore in order to ensure that adequate glycosylation takes place in accordance to the invention, the skilled person will seek possible mutations for the antibody fragment, which results in the availability of an additional acceptor group which can act as a glycosylation site, said acceptor group being located before the first hydrophobic stretch, and said acceptor group in the 3D structure being located below plane (2) of figure 2.

It is believed to be well within the ability of the skilled person based on the above teachings to select the appropriate mutation sites in the antibody fragments.

For example it can be expected that the introduction of a glycosylation site at amino acid 11 (3) or the introduction of a glycosylation site at amino acid 17 (4) will indeed lead to a better secretion due to the fact that the glycosylation site is located before the hydrophobic stretch and also substantially below plane (2).

Glycosylation at amino acids 60 (5), 81 (6) or 82b(7) are borderline cases, because although in the linear sequence the location of the glycosylation groups would be indeed between the N-terminus and the hydrophobic stretch, figure 2 clearly illustrates that these amino acids are located fairly close to plane (2) and hence the glycosylation groups may or may not interact with the CDR regions, dependant on the orientation of the glycosylation groups.

Similarly for ScFv fragments, being associated complexes of a V_H and a V_L domain (production see L.G.J. Frenken et al (1994)). The structure is believed to be remarkably similar to the stucture of R6 of Figure 2, however now a V_L fragment is associatively bonded to the V_H domain around amino acid number 60 (numbering in accordance to Kabatt).

The amino acid sequence of the Vh fragment is for example given in Example III. The amino acid sequence of the light chain is as follows:

```
10        20        30        40        50
     ....|....|....|....|....|....|....|....|....|....|
D1.3 DIELTQSPASLSASVGETVTITCRASGNIHNYLAWYQQKQGKSPQLLVYY

                60        70        80        90        100
     ....|....|....|....|....|....|....|....|....|....|....|
D1.3 TTTLADGVPSRFSGSGSGTQYSLKINSLQPEDFGSYYCQHFWSTPRTFGG

     ....|...
D13  GTKLEIKR
```

The sequence of a possible peptide linker is as follows:

GGGGSGGGGSGGGGS

Based on this structure and the considerations as explained above, the skilled person would expect that the introduction of a glycosylation site in the V_H part by mutation of the 11 position (L11N, A13S, P14A) or a mutation at the 17 position (S17N) would lead to improved secretion.

Based on the above-presented reasoning glycosilation at amino acid positions 108(8) (T108N) or 110(9) (T110N) in the V_H part, however is likely not to result in the desired effect.

Glycosylation at amino acids 81 (K81N and N82aS) or 82b(S82bN) or 5 (Q5N) in the V_H part are borderline cases, because, although in the linear sequence the location of the glycosylation groups would be indeed between the N-terminus and the hydrophobic stretch, these amino acids are most likely located fairly close to the CDR area and hence the glycosylation groups may or may not interact with the CDR regions, dependant on the orientation of the glycosylation groups.

Additionally the following mutations in the V_L part are believed to be successful: at the 5 position T5N, at the 10 position S10N, at the 12 position S12N, at the 18 position T18N.

Applicants believe that based on the above teaching the skilled persons will be able to improve the secretion of other relatively hydrophobic proteins. Especially good results are expected for the production of antibodies or antibody fragments, in particular antibodies which are devoid of light chains or fragments thereof.

The invention also involves the provision of novel glycosylated proteins comprising at least one glycosylation group and at least one hydrophobic stretch, said glycosylation group being located between the N-terminus of the protein and the hydrophobic stretch, and wherein in the 3 dimensional form of the protein,

1. the distance between the glycosylation site and the binding site(s) and/or active site(s) is greater than the length of the glycosylation group; or

(b) the distance between the glycosylation site and the binding site(s) and/or active site(s) is less than the length of the glycosylation group, provided the orientation of the glycosylation group is such that it cannot interact with the binding site(s) and/or the active site(s) of the protein.

The invention further comprises gene sequences capable of encoding this protein. Preferably said gene sequence comprising at least one partial sequence encoding a hydrophobic stretch in the protein and a second partial sequence encoding a glycosylation site located between the N-terminus of the protein and the hydrophobic stretch. The

invention also relates to vectors comprising such a gene sequence as well as a genetically modified eukaryotic cell, comprising capable of producing the proteins of the invention or comprising this vector.

Furthermore the invention relates to consumer products, for example food products or products for cleaning or treatment of fabrics, the human body or hard surface, said consumer products comprising a protein in accordance to the invention.

The invention will be further illustrated in the examples.

Example I

Wild type (CY000), and mutant forms of *Fusarium solani pisi* cutinase were expressed in *Saccharomyces cerevisiae* (strain CEN. PK 111-32D) using the vectors and methods as described in Sagt CM et al, (1998) and in van Gemeren I. Et al (1995). The mutations were introduced by Polymerase Chain Reaction (PCR) as described in Hedstrom L, et al (1991). Standerd techniques of molecular cloning were used as described in Sambrook, J., Fritsch, E.F. and Maniatis (1989).

A mutant cutinase was produced (CY047) with an A29S mutation. This mutation introduces a glycosylation site into the cutinase.

A mutant was produced (CY028) with G82A, A85F, V184I, A185L, L189F mutations (as described in WO 94/14963) This modification results in a significantly increased hydrophobicity and increased binding affinity with respect to the binding protein (BiP) and hence a significantly higher BiP score. The aggregated BiP score of said protein is about 80 units higher than the BiP score of the corresponding wild-type.

A mutant (CY181) was produced with G82A, A85F, V184I, A185L, L189F mutations as well as a A29S mutation. These mutations introduce both a glycosylation site and hydrophobicity into the protein. The glycosylation site is located between the N-terminus of the cutinase and the

first hydrophobic stretch. Furthermore, the glycosylation site is located in the 3D structure at the lower end as illustrated in figure 1.

A mutant (CY182) was produced with G82A, A85F, V184I, A185L, L189F mutations as well as a R211N mutation. In this mutant the glycosylation site is located between the C-terminus of the cutinase and the hydrophobic stretch.

S. cerevisiae transformants expressing wild-type cutinase and the mutants were each grown overnight in YP 2% glucose (YP = 1% Difco yeast extract, 2% bacto peptone)and diluted 1:10 in YP 5% galacotose to induce cutinase production. After 24 hours, samples of 1 ml were taken and centrifuged for 1 min at 14,000 X g to pellet the yeast cells. The amount of secreted cutinase in the medium was determined with the para-nitopheny1 butyrate (PNPB, Sigma) assay as described in Kolattukudy et al (1981). The results are shown in figure 5.

The wild-type and the CY047, CY028 and CY181 mutant cutinase expressing cells were also grown under the same conditions, whereby the 5% galactose solution also contained 5 µg/ml tunicamycin to inhibit glycosylation. The results are shown in figure 6.

These results clearly show that glycosylation of the Wild type cutinase does not materially affect the amount of secreted cutinase. The introduction of hydrophobicity, however, greatly reduced the secretion of cutinase (compare Wild-type and CY028). The introduction of glycosylation,

however significantly restored the secretion efficiency (compare CY028 and CY181). The reason for this restored efficiency is indeed the presence of glycosylation before the hydrophobic stretch. This is clearly shown in the experiments where glycosylation is inhibited by tunicamycin leading to a significant reduction of secretion efficiency for the CY181. It can be concluded that CY182 cutinase with the glycosylation site at the C-terminus is not as efficiently secreted as CY181. This indicates that the site of glycosylation is more powerful in enhancing the secretion if it is located before the hydrophobic stretches than when it is located after the hydrophobic stretches. This could be due to the decreased affinity for BiP of CY181.

Example II

Construction of *Pichia pastoris* strains carrying cutinase genes and the extracellular production of cutinase and cutinase mutants by this lower eukaryote.

The *Pichia pastoris* strain GS115 (*his4*, Mut^+, Invitrogen, USA) was used. Cutinase variant genes were excised from the *S. cerevisiae* expression plasmids described above by digestion of the plasmid DNA with *Sac*I and *Hin*dIII. The cutinase genes were ligated into the *Sac*I/*Hin*dIII vector fragment derived from a modified pBR322 (Clonetec, USA) vector, that contained a linker with a *Bam*HI, *Sac*I, *Hin*dIII and an *Eco*RI site in that order inserted at between the *Bam*HI and *Eco*RI sites in pBR322. The cutinase encoding fragment released from this intermediate

vector by digestion with *Bam*Hl and *Eco*R1 was ligated into the expression vector fragment of pPIC9 after digestion with *Bam*HI and *Eco*RI. This placed the cutinase encoding sequences, linked to the *SUC*2 secretion signal sequence under the control of the *AOX* promoter of pPIC9. Before transformation, the plasmid was linearized with *Aat*II.

For transformation of *Pichia pastoris, Pichia pastoris* GS115, cells were grown overnight in YP 2% glucose in shake flasks at 30°C. 0.1-0.5 ml of the overnight culture was inoculated in 100 ml of fresh YP 2% glucose (to an OD_{600} of approximately 0,3) and grown o/n at 30°C until the OD_{600} reached 1,3-1,5. The cells were centrifuged at 4000 × g for 5 min. at 4°C and washed twice in 100 and 50 ml ice-cold sterile water and in 4 ml ice-cold 1M sorbitol. After centrifuging (4000 × g, 5 min., 4°C) the pellet was resuspended in 1 ml ice-cold 1M sorbitol. 75 µl of cells were mixed with 12.75 µg of linearized DNA respectively and transferred to an ice-cold electroporation cuvette (*E. coli* pulser cuvette, Bio-Rad) and the cuvette with the cells was incubated on ice for 5 min. The cuvette was transferred to a Bio-Rad gene pulser and the cells were transformed by application of a pulse of 1.5 kV, 25 µF and 400 Ω. After the pulse the cells were immediately transferred into 800 µl YP 2% glucose at 30°C and were incubated for one hour at 30°C. The cells were washed in 1M sorbitol and thereafter resuspended in 200 µl sorbitol. The cells were plated on YNB 1% glucose plates (YNB= Yeast Nitrogen Base without amino acids (a.a.), Difco), and the plates were incubated at 30°C.

P. pastoris transformants containing expression cassettes for either CY000, CY028 or CY181 cutinase mutants (see example I) that were able to grow on YNB 1% glucose plates after electroporation were inoculated in 10 ml of BMG (100mM potassium phosphate buffer, pH 6.0, 1X YNB without a.a., 1% glycerol) in a 400 ml shake flask and grown at 30°C at 200 rpm until the culture reached an OD_{600} of 2-6. The cells were harvested by centrifugation at 4000 X g for 5 min. at room temperature. The supernatant was decanted and the cell pellet was resuspended in 50 ml BMMY (100mM potassium phosphate buffer, pH 6.0, 1X YNB w/o a.a., 1% Difco yeast extract, 2% bacto peptone 0.5% methanol) to induce expression of the cutinase variants. Cells were grown in a 400 ml shake flask at 30°C at 200 rpm. To maintain induction, 250 µl of 100% Methanol was added every 24 hours. Every 24 hours, just before adding the methanol, 1 ml of sample was taken. The sample was centrifuged for 1 min. at 14000 X g and extracellular cutinase levels were determined by activity assays with p-nitrophenyl butyrate as substrate (PNPB, Sigma)as described above. The cells and the supernatants were stored at -20°C for further analysis.

In order to compare the differences in secretion between the CY000, CY028 and CY181 mutant cutinases, the maximum amount of extracellular cutinase obtained at the end of log phase growth was divided by the OD_{600} at that time-point. This results in the maximum cutinase secretion in mg/l per OD_{600}. 4 different CY000 Mut^+ and 4 different CY000 Mut^S transformants were studied to obtain these values. The same was done for the CY028 samples except that in this case only 3 different CY181

Mut$^+$ and 4 different CY181 Mut^S transformants were analysed. There is little difference in secretion between Mut$^+$ and Mut^S strains; the maximal amount of secreted cutinase did not differ significantly between Mut$^+$ and Mut^S.

The average relative secretion of CY028 and CY181 cutinase compare to CY000 secretion was calculated. When the secretion of CY000 is set at 100%, CY028 is secreted at 41% and CY181 is secreted at 79%. This again shows that, compared to CY000, the hydrophobic transformant CY028 is thus significantly less well secreted. However, secretion can be restored to an acceptable level by introducing of the glycosylation group in CY181.

A *P. pastoris* CY181 transformant was also grown under the same conditions, but with the addition of 10µg/ml tunicamycin in the BMMY medium to prevent glycosylation. This lead to secretion that was significantly lower than the secretion of CY181 grown in the absence of tumicamycin and more similar to the level of secretion of CY028.

Example III

Cloning of the variable domain of lama antibodies (HC-V) raised against the azo-dye RR6 and improved secretion of these HC-V's by *S. cerevisiae*

This example refers to the production of single domain heavy chain antibody fragments where improved secretion yield can be achieved in yeast by genetic mutation at

specified positions in the antibody fragment genes so as to introduce amino acids that are capable of acting as glycosylation sites.

a) Induction of humoral immune responses in llama.

Male llamas were immunized with a water in oil emulsion (1:9 V/V, antigen in water: specol, subcutaneously and intramuscularly. 300 µg of the azo-dye Reactive Red-6 antigen (Procion Rubine MX-B, ICI) coupled to bovine serum albumin via its reactive triazine group, were inoculated per immunization site in 0.75-1.5 ml water / oil emulsion.

Further immunizations were performed at 3 weeks and 5 weeks from the date of initial immunization.

b) Isolation of gene fragments encoding llama HC-V domains.

A 200 ml blood sample was then taken from the inoculated llamas and peripheral blood cells were obtained via a Ficoll-Paque centrifugation (Pharmacia). The total RNA was isolated from about 10^7 lymphocytes essentially as described by Chomczynnski and Sacchi (1987). Following cDNA synthesis (with the Amersham first strand cDNA kit, Amsersham-Pharmacia, UK), DNA fragments encoding HC-V fragments and part of the long or short hinge region were amplified by a PCR using the specific primers:

*Pst*I

V_H - 2B 5'-AGGTSMAR**CTGCAG**SAGTCWGG-3'

S = C and G, M = A and C, R = A and G, W = A and T,

Lam-07 *Hin*dIII

5'-AACAGTT**AAGCTT**CCGCTTGCGGCCGCGGAGCTGGGGTCTTCGCTGTGGTGCG-3'

Lam-08 *Hin*dIII

5'-AACAGTT**AAGCTT**CCGCTTGCGGCCGCTGGTTGTGGTTTTGGTGTCTTGGGTT-3'

The PCR reaction was carried out for 32 cycles of 1 min 94 °C, 1.5 min at 55 °C and 2 min at 72 °C.

Upon digestion of the PCR fragments with *Pst*I and *Bst*EII, the DNA fragments with a length between about 300 and 450 base pairs encoding the HC-V domain, but lacking the first three and the last three codons, were purified via gel electrophoresis (using a Quiaex DNA isolation kit, Quiagen).

c) Construction of *S. cerevisiae* episomal expression plasmids encoding llama HC-V domains.

Saccharomyces cerevisiae episomal expression plasmids pUR4547 and pUR4548 were derived from pSY1 (Harmsen et al., 1993) and have been deposited at the Centraal Bureau voor Schimmelcultures, Baarn deposition numbers: CBS 100012 and CBS 100013, respectively. See restriction maps of figures 3 and 4.

Both plasmids contain the *GAL7* promoter and *PGK* transcription terminator sequences as well as the invertase (*SUC2*) signal sequence. The signal sequence is followed by the first 5 codons of the HC-V domain (including the *Pst*I site), a stuffer sequence which can be removed by digestion with *Pst*I/*Bst*EII, then the last six codons of the HC-V domain.

In pUR4547, these six codons are followed by two stop codons, an *Afl*II and a *Hin*dIII site. In pUR4548, these six codons are followed by eleven codons encoding a *myc* epitope tag, two stop codons, an *Afl*II and *Hin*dIII site.

A vector fragment was obtained from pUR4548 by digestion with *Pst*I and *Bst*EII restriction endonulceases, and ligated with the approximately 300-450 bp *Pst*I - *Bst*EII fragments obtained as described in (b.) The effect of this is to create a genetic library where fragments encoding particular HC-V domains are contained within a single plasmid.

After transformation of *S. cerevisiae* (Finlayson S,D., et al 1991) transformants were selected from minimal media agar plates (comprising 0.7 % Difco yeast nitrogen base, 2 % glucose and 2 % agar).

d) Screening for antigen specific HC-V domains.

For the expression and secretion of llama HC-V fragments linked to the myc-tag, individual transformants containing

plasmids derived from pUR4548 were grown overnight in selective minimal medium (0.7 % yeast nitrogen base, 2 % glucose) and subsequently diluted ten times in YPGal medium (1 % yeast extract, 2 % bacto peptone and 5 % galactose).

After 24 and 48 hours of growth, the culture supernatants of the transformants were analysed by enzyme linked immunosorbent assay (ELISA) for the presence of HC-V fragments that specifically bind to the antigen RR6. For the ELISA, RR6 was covalently coupled to covalink plates (NUNC, 478042) and after incubation with the culture supernatants, bound antibodies were detected using the anti-myc monoclonal antibody 9E10 (ATCC) and a polyclonal goat-anti-mouse-HRP conjugate (Bio-Rad, 172-1011).

In this way, a number of anti-RR6 HC-V fragments were isolated, among which are those referred to as R2, R5 and R7. These fragments have the following amino acid sequences.

```
                10        20        30            40
        ....|....|....|....|....|....|.abc...|....|....|..
R2      QVQLQESGGGLVQAGGSLRLSCAASGRATSGHGHYGMGWFRQVPGKEREF
R5      QVQLQESGGGLVQAGGSLRLSCAASGRTSHGYGGYGMGWFRQVPGKEREL
R7      QVQLQESGGGLVQTGDSLRLSCAASGRTSHGYGGYGMGWFRQIPGKEREL

          50        60        70        80            90
        ..|..a..|....|....|....|....|....|..abc..|....|...
R2      VAAIRWSGKETWYKDSVKGRFTISRDNAKTTVYLQMNSLKPEDTAVYYCA
R5      VAAIRWSGTSTYYADSVKGRFTISRDNVKNMVYLQMNSLKPEDTAVYHCA
R7      VAAIRWSGRNTYYADSVKGRFTISRDNVKDMLYLQMDSLKPEDTAVTYCA


            1.                  110
        .|....|abcdefgh....|....|...
R2      ARPVRVDDISLPVGFDYWGQGTQVTVSS
R5      ARTVRVVDISSPVGFAYWGQGTQVTVSS
R7      VRTVRVVDISSPVGFAYWGQGTQVTVSS
```

The gene fragments encoding these anti-RR6 fragments were subcloned from the pUR4548 into pUR4547 resulting in the plasmids shown in table 1:

Table 1: plasmids

Fragment	pUR4548 derivative	pUR4547 derivative
R2	pUR4633	pUR4643
R5	pUR4636	PUR5353
R7	pUR4638	pUR4644

Construction of *S. cerevisiae* multicopy integration vectors for the expression of llama HC-V domains combines the benefits of high copy number and mitotically stable expression. The concept of a multicopy integration system into the rDNA locus of lower eukaryotes and the proven

procedures for obtaining such constructs are described by Giuseppin *et al.* 1991 (WO-A-91/00920).

e) Construction of glycosylation mutants of HC-V(RR6) with improved secretion yield

Of the anti-RR6 HC-V fragments identified in (d), R2, R5 and R7 were subjected to genetic modifications to introduce amino acids that predispose the protein to glycosylation. R2 encoded by the plasmids pUR4643 and pUR4633 is efficiently secreted by *S. cerevisiae* but the total production level is low. R5, is well produced but accumulates intracellularly if expressed in *S. cerevisiae* from the plasmids pUR4636 or pUR5353. R7 is produced at a relatively high level in *S. cerevisiae* from the plasmids pUR4638 or pUR4644 but like R5, the majority of the protein remains intracellular.

In total, six different replacements are described for the three different antibody fragments. All mutations were performed by making use of the overlap-extension PCR technique (Horton *et al.*, 1989). The Amino acid replacements, the corresponding nucleotide sequences of the oligo's used to introduce these mutations and the oligo' names are shown in the table 3.

The plasmids pUR4643 (R2 fragment), pUR4636 (R5 fragment for mutations 1, 2, 5, 6 and 7) pUR5353 (R5 fragment for mutation 3) and pUR4644 (R7 fragment) were used as

templates for the DNA modifications in the HC-V fragment coding seguences.

The technique of splicing by overlap PCR extension required the creation of separate left hand and right hand PCR fragments and the subsequent combination of the two to generate a coding region for the desired mutations.

Left-hand (=Upstream) fragments were created by making use of the 5′ oligo′ JA-L-gl (table 2) and depending on the desired modification, one of the 3′ oligos designated with an 'L' shown in table 2. The first number in the oligo′ corresponds to the location of the amino acid replacement and the last number, or numbers, to the antibody fragment to be modified. The right-hand (=downstream) fragments were created by making use of the 3′ oligo JA-R-gl (table 2) and the 5′ oligo′ containing the inverse complement of part of the 3′ 'L' oligo.

Table 2. The 5′ and 3′ outermost PCR primers

Oligo′ Number	Oligo′ Sequence
JA-R-gl	TAGCTCACTCATTAGGCACC
JA-L-gl	GCCTTTAGCTATGTTCAG

PCR was performed using the following conditions: 1 min at 95 °C, 1.5 min at 50 °C and 1 min at 72 °C for 25 cycles. Subsequently, the resulting PCR fragments were combined by performing a PCR reaction with the appropriate left- and right-hand fragments as templates and JA-L-gl and JA-R-gl

oligo's as primers. The conditions used were as described above.

The products of these reactions were digested either with *Pst*I and *BstE*II (R2 and R5 mutants 1, 2, 5, 6 and 7 and R7 mutants 1, 2 and 3) or *Pst*I and *Hin*dIII (R2 mutant 3 and R5 mutant 3 in which the *Bst*EII site is removed by the mutation) and the HC-V coding fragment isolated and cloned into the pUR4547 vector fragment obtained by digestion with *Pst*I/*BstE*II or *Pst*I/*Hin*dIII as appropriate.

The expression plasmids so formed are listed in table 3.

Saccheromyces cerevisiae was transformed with these plasmids and expression of the HC-Vs induced as described above. Samples of the culture supernatants from these transformants were analysed by polyacrylamide gel electrophoresis (PAGE) using a Multipho II (Amersham-Pharmacia) or Mini-Protean (Bio-Rad) system according to the manufacturers instructions. As glycosylated proteins do not bind Coomassie Brilliant Blue protein stain as efficiently as non-glycosylated proteins, samples were also treated with Endo H (Boehringher) according to the manufacturers instructions to remove the glycosylation.

The presence of the glycosylation sites had a marked effect on the antibody fragments produced. The results for R2 showed that the secreted forms of the modified HC-V fragments were indeed glycosylated as could be seen by an increase in the apparent molecular weight of the proteins

which restored to normal after Endo H treatment. Mutants number 1 and 3 appear only to be core glycosylated as they show a more discrete band on the PAGE gels compared to the other variants which show diffuse higher molecular weight bands. Glycosylation had no effect on the levels of secretion of the R2 derivatives. However, for R5 and R7 (the poorly secreted proteins) significant improvements in the levels of secretion were observed for the mutant R5-1, R5-2, and for the R7 variants R7-1 and R7-2. The mutant 3 derivatives showed no improvement in secretion. This is in accordance with the invention as R5-1, R5-2 and R7-1 and R7-2 comprise glycosilation sites before the hydrophobic stretch.

Slight improvement was found for mutants R5-6 and R5-7.

These results are summarised in table 4.

Table 4 Results of antibody glycosilation experiments

RR6 binding fragment	Mutation number	Secretion
R2	1	+
	2	+
	3	+
	5	+
	6	+
	7	+
R5	1	+
	2	+
	3	-
	6	slight
	7	slight
R7	1	+
	2	+
	3	-

References

Blond-Elguindi, S. et al (1993) Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specifity of Binding Protein (BiP) Cell 75: 717-728

Chomczynnski, P. and Sacchi, N. (1987) Single step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. Analytical Biochem. 162, 156-159.

Finlayson S.D., Fleming C., Berry D.R. and Johnston J. An improved lithium-acetate method for yeast transformation Biotechnology Techniques (1991) 5:13-18.

Frenken L.G.J. et al in Nato ASI Series, Vol. H 82, Biological Membranes: Structure, Biogenesis and Dynamics, Springer Verlag Berlin Heidelberg, 1994, pages 223-236.

Harmsen, M.M., Langedijk, A.C., van Tuinen, E., Geerse, R.H., Raue, H.A., and Maat, J. (1993) Effect of a *pmr 1* disruption and different signal sequences on the intracellular processing and secretion of *Cyamopsis tetragonoloba* alpha-galactosidase by *Saccharomyces cerevisiae*. Gene 125: 115-123.

Hedstrom L, Graf L, Stewart CB, Rutter WJ, Phillips MA Modulation of enzyme specificity by site-directed mutagenesis. Methods Enzymol 1991; 202:671-87.

Horton R.M., Hunt H.D., Ho S.N., Pullen J.K. and Pease L.R. Engineering Hybrid Genes Without The Use Of Restriction Enzymes - Gene-Splicing By Overlap Extension Gene 77: 61-68 1989

Kolattukudy, P.E., Purdey, R.E. and Maiti I.B. (1981) Cutinases from fungi and pollen. Methods in enzymology 71: 652-664.

Kurucz I. Et al, Proc. Nat. Acad. Sci. USA 90:3830-3834 (1993).

Sagt C.M.J, Muller WH, Boonstra J, Verkleij AJ, Verrips CT Applied and Environmental Microbiology, Jan 1998 64(1):316-24 64(1):316-32.

Sambrook, J., Fritsch, E.F. and Maniatis (1989) Molecular cloning: a laboratory manual, 2 nd edition. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, N.Y.

van Gemeren IA, Musters W, van den Hondel CA, Verrips CT Construction and heterologous expression of a synthetic copy of the cutinase cDNA from *Fusarium solani pisi*. J Biotechnol 1995; 40:155-62.

RR6 binding fragment	Mutation Number	Mutation	Olgo' Number.	Olgo' sequence	Plasmid
R2	1	S82bN, K83T, P84A	**JA01R2** **JA01L2**	GCAAATGAACAACCTGACAGCTGAAGATACGG CCGTATCTTCAGCTGTCAGGTTGTTCATTTGC	pUR5336
	2	L11N, Q13T	**JA02R25** **JA02L25**	GGGGGAGGAAATGTGACTGCTGGGGGC GCCCCCAGCAGTCACATTTCCTCCCCCC	pUR5337
	3	T110N	**JA03R2567** **JAV03L2567**	GGACCCAGGTCAACGTCTCCTC GAGGAGACGTTGACCTGGGTCC	pUR5338
	5	Q81N, N82aT	**JA05R2** **JA05L2**	JAGGTTTATCTGAACATGACCAGCCTGAAACC GGTTTCAGGCTGGTCATGTTCAGATAAACC	pUR5340
	6	K60N	**JA06R2** **JA06L2**	AGACATGGTATAATGACTCCGTGAAGG CCTTCACGGAGTCATTATACCATGTCT	pUR5341
	7	S17N, R19S:	**JA07R256** **JA07L256**	CTGGGGCAATCTGTCACTCTCCTGTGC GCACAGGAGAGTGACAGATTGCCCCCAG	pUR5342
R5	1	S82bN, K83T, P84A	**JA01R5** **JA01L5**	GCAAATGAACAACCTGACAGCTGAGGACACGG CCGTGTCCTCAGCTGTCAGGTTGTTCATTTGC	pUR5343
	2	L11N, Q13T	**JA02R25** **JA02L25**	GGGGGAGGAAATGTGACTGCTGGGGGC GCCCCCAGCAGTCACATTTCCTCCCCC	pUR5344
	3	T110N	**JA03R2567** **JA03L2567**	GGACCCAGGTCAACGTCTCCTC GAGGAGACGTTGACCTGGGTCC	pUR5345
	6	K60N	**JA06R56** **JA06L56**	GTACATACTATAATGACTCCGTGAAGG CCTTCACGGAGTCATTATAGTATGTAC	pUR5348
	7	S17N, R19S	**JA07R256** **JA0L256**	CTGGGGGCAATCTGTCACTCTCCTGTGC GCACAGGAGAGTGACAGATTGCCCCCCAG	pUR5349
R7	1	S82bN, K83T, P84A	**JA01R7** **JA01L7**	GCAAATGAACAACTTGACAGCTGAGGACACGG CCGTGTCCTCAGCTGTCAAGTTGTTCATTTGC	pUR5350
	2	L11N, Q13T	**JA02R7** **JA02L7**	GGGGGAGGAAATGTGACTGCTGGGGAC GTCCCCAGCAGTCACATTTCCTCCCCC	pUR5351
	3	T110N	**JA03R2567** **JAV03L2567**	GGACCCAGGTCAACGTCTCCTC GAGGAGACGTTGACCTGGGTCC	pUR5352

Table 3

Claims

1. A method for producing a protein, preferably a heterologous protein, in a lower eukaryotic cell, followed by secretion of said protein from said cell, said protein comprising at least one hydrophobic stretch, wherein the protein is glycosylated such as to increase the level of secretion of the protein, without substantially affecting the functionality of the protein and wherein at least one N-glycosylation group is located between the N-terminus of the protein and the first hydrophobic stretch and with the proviso that

a) the protein is not a single chain Fv molecule against human TfR with an N linked glycosilation site at position 19 in FR1 of V_H, and

b) the protein is not a single chain Fv molecule against hapten DNP with an N linked glycosilation site at position 19 in FR1 of V_H.

2. A method according to claim 1, wherein the hydrophobic stretch of the protein has an individual BiP score of 10 or more and/or the protein has an aggregated BiP score which is at least 50 units above the BiP score of the corresponding wild-type protein.

3. A method according to one or more of the preceding claims wherein in the 3 dimensional form of the protein,

(a) the distance between the glycosylation site and the binding site(s) and/or active site(s) is greater than the length of the glycosylation group; or
(b) the distance between the glycosylation site and the binding site(s) and/or active site(s) is less than the length of the glycosylation group, provided the orientation of the glycosylation group is such that it cannot interact with the binding site(s) and/or the active site(s) of the protein.

4. A method according to one or more of the preceding claims wherein the protein has a substantially ellipsoidal three dimensional shape, wherein the binding and/or active site(s) are located at the upper end of the ellipsoid and the glycosylation groups are located at a substantial distance from said upper end.

5. A method according to one or more of the preceding claims, wherein the protein is a lysozyme, cutinase or antibodies, especially variable domains thereof.

6. A method according to any of the preceding claims, wherein the eukaryotic cell is a yeast or a mould.

7. A method according to claim 6, wherein the yeast belongs to the genera Saccharomyces, Kluyveromyces, Hansenula or Pichia.

8. A method according to claim 6, wherein the mould belongs to the genera Aspergillus or Trichoderma.

9. Glycosylated protein comprising at least one glycosylation group and at least one hydrophobic stretch, said glycosylation group being located between the N-terminus of the protein and the hydrophobic stretch, and wherein in the 3 dimensional form of the protein,
(a) the distance between the glycosylation site and the binding site(s) and/or active site(s) is greater than the length of the glycosylation group; or
(b) the distance between the glycosylation site and the binding site(s) and/or active site(s) is less than the length of the glycosylation group, provided the orientation of the glycosylation group is such that it cannot interact with the binding site(s) and/or the active site(s) of the protein.

10. A gene sequence capable of encoding the glycosylated protein of claim 9 or the protein obtainable in the process according to one or more of claims 1-8.

11. A gene sequence according to claim 10, comprising at least one partial sequence encoding a hydrophobic stretch in the protein and a second partial sequence encoding a glycosylation site located between the N-terminus of the protein and the hydrophobic stretch.

12. A vector comprising a gene sequence in accordance to claim 10 or 11.

13. Genetically modified eukaryotic cell, capable of producing a protein in accordance to claim 9 or comprising a vector according to claim 12

14. Consumer product comprising a protein in accordance to claim 9 or a gene according to claim 10 or a genetically modified eukaryotic cell according to claim 13.

Fig.1.

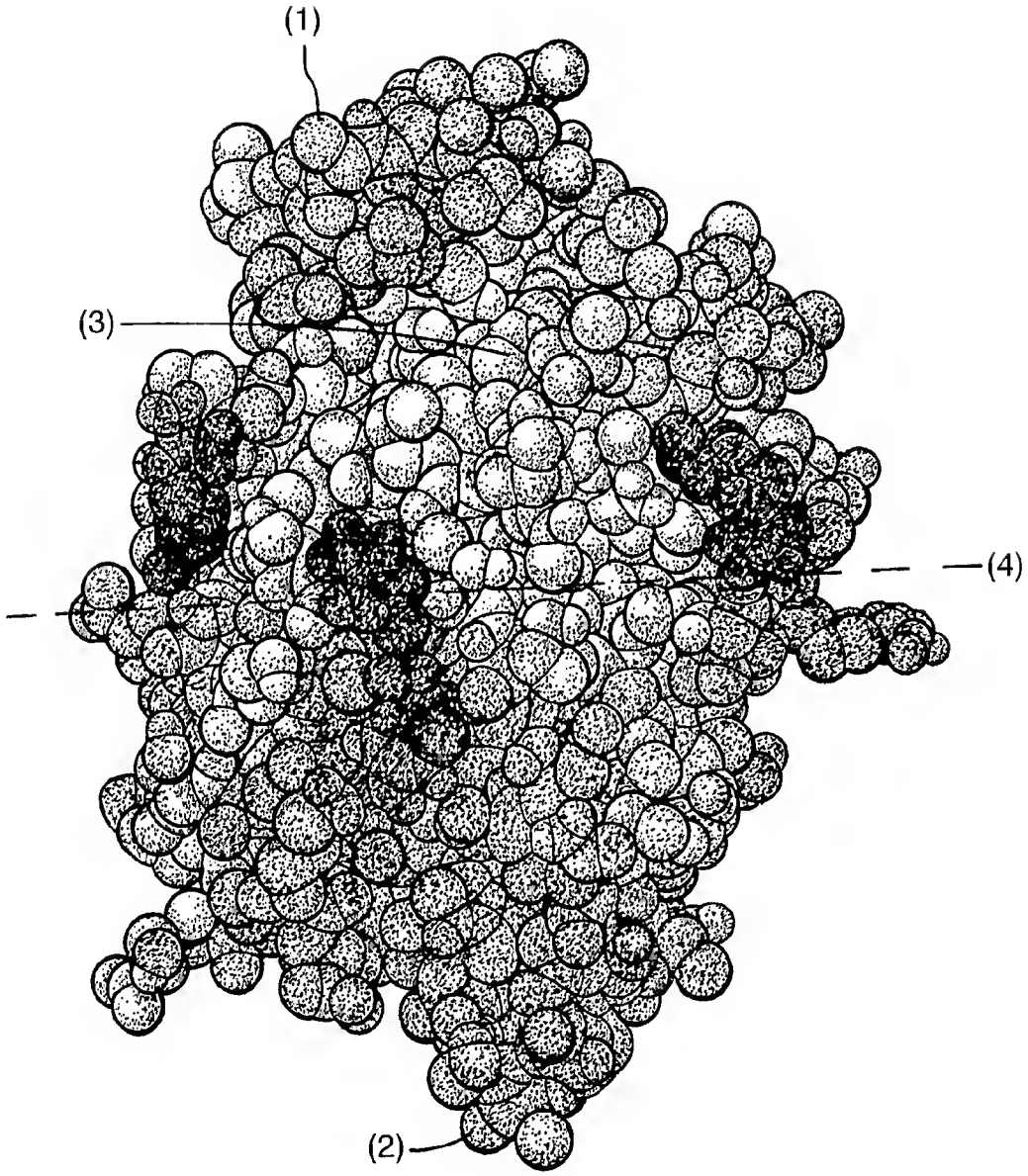


Fig.2.

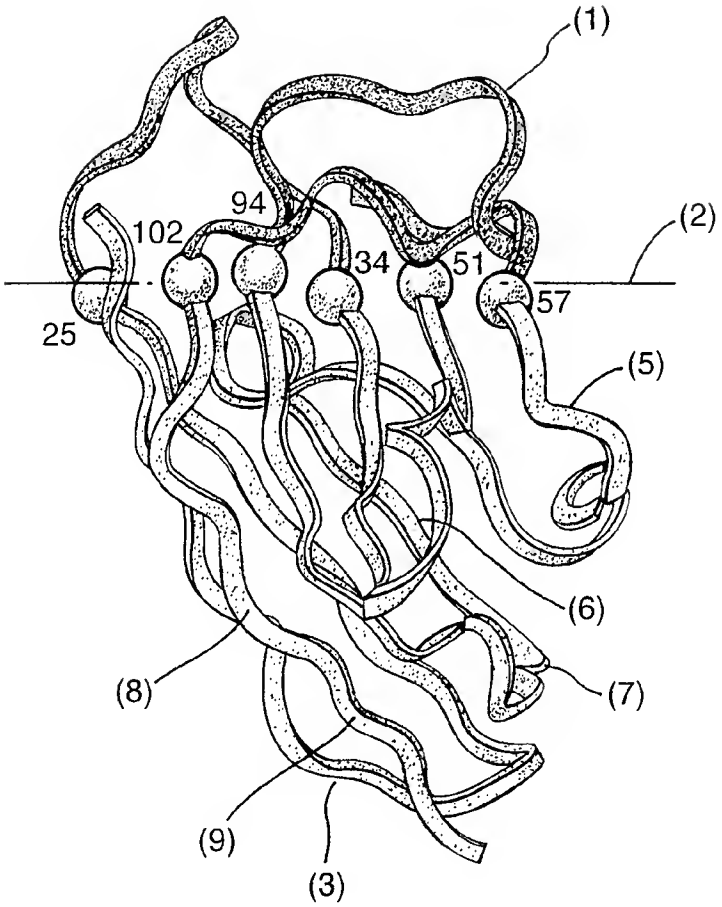
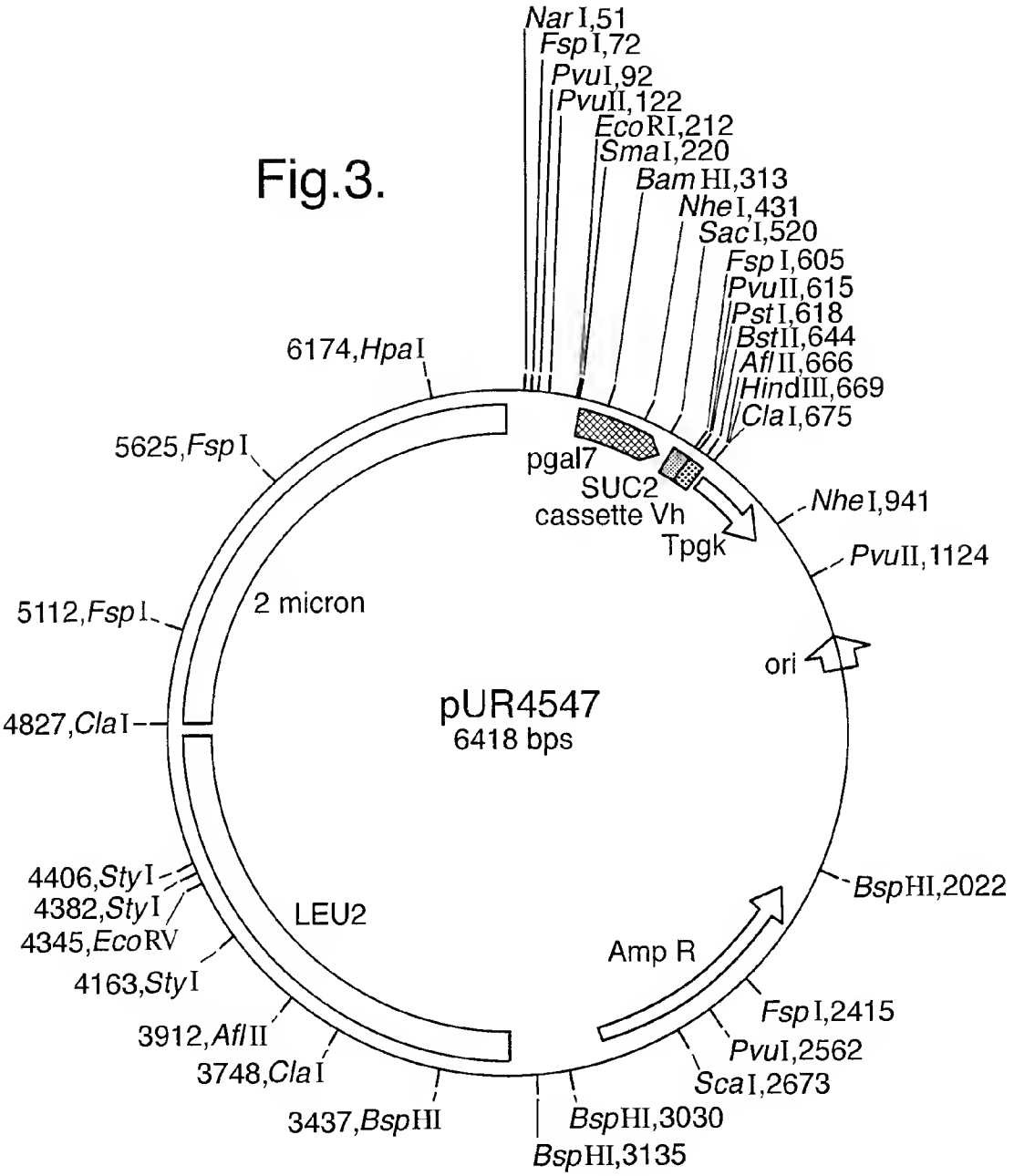
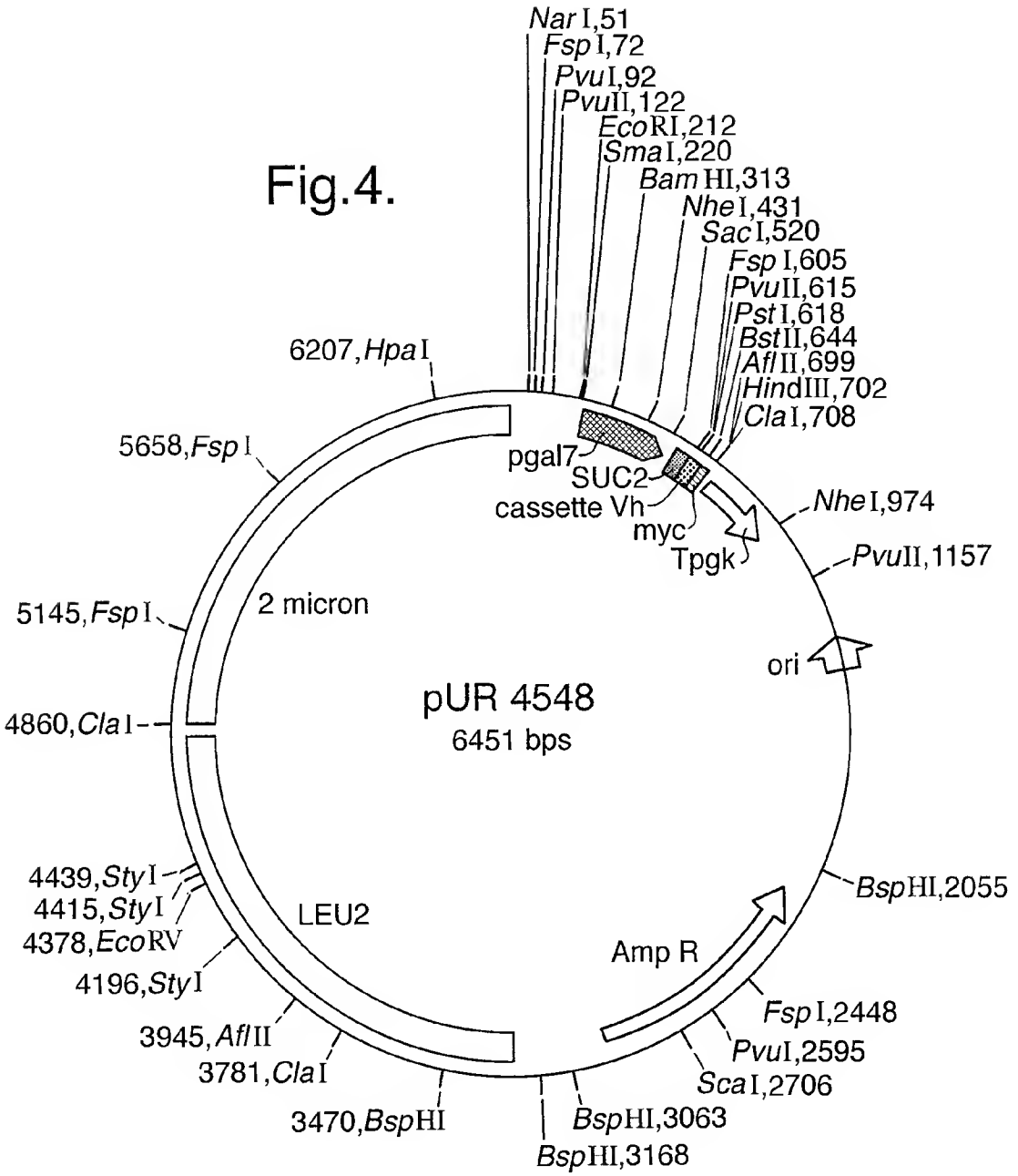


Fig.3.




```
                                                    AflII
PstI                          BstEII                HindIII
     <-                    MPG 124 RL (47)          ->
CTGCAGGAGTCATAATGAGGGACCCAGGTCACCGTCTCCTCATAATGACTTAAGCTT
--------+---------+---------+---------+---------+-------  57
GACGTCCTCAGTATTACTCCCTGGGTCCAGTGGCAGAGGAGTATTACTGAATTCGAA
 <-                    MPG 125 RL (55)                ->
L  Q  E  S  *  *  G  T  Q  V  T  V  S  S  *  *  L  K  L
```

Fig.4.




```
                                                                    AflII
PstI                          BstEII                                HindIII
     <-     MPG 126 RL (31)       ->-<-              MPG 128 RL (49)       ->
CTGCAGGAGTCATAATGAGGGACCCAGGTCACCGTCTCCTCAGAACAAAAACTCATCTCAGAAGAGGATCTGAATTAATGACTTAAGCTT
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 90
GACGTCCTCAGTATTACTCCCTGGGTCCAGTGGCAGAGGAGTCTTGTTTTTGAGTAGAGTCTTCTCCTAGACTTAATTACTGAATTCGAA
 <-          MPG 127 RL (45)              ->-<-          MPG 129 RL (43)            ->
L  Q  E  S  *  *  G  T  Q  V  T  V  S  S  E  Q  K  L  I  S  E  E  D  L  N  *  *  L  K  L
                                  <-        myc tail'          ->
```

Fig.5.

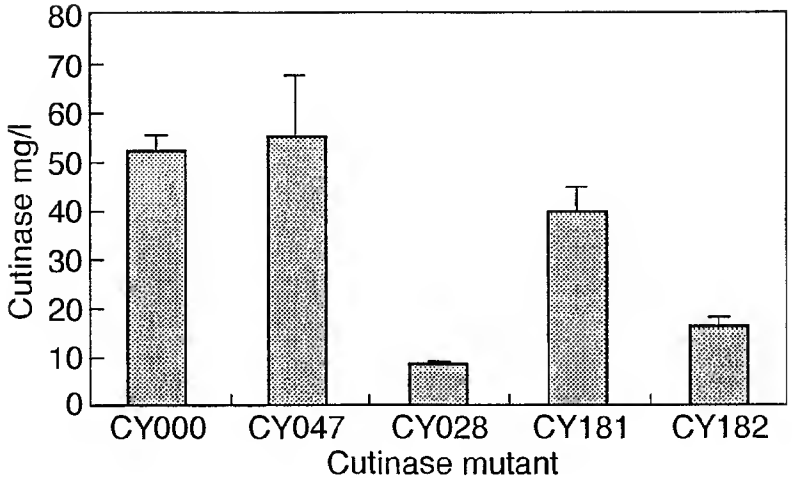


Fig.6.